

CLAIMS

1. An information system that includes a plurality of application servers, each of which can execute at least one application program for processing a transaction received  
5 from a terminal, and a database server for accessing a database based on a request received from any one of the plurality of application servers, comprising:

a process time monitoring unit, for monitoring processing time required for the application program to  
10 process a transaction received by an application server;

a bottleneck identification unit, for identifying, based on monitoring results of processing time, a bottleneck in at least one of the plurality of application servers if the processing time is not within a  
15 predesignated permissible range; and

a load controller, for reducing a multiplicity of the application program on an application server identified as having a bottleneck.

2. An information system according to claim 1, wherein  
20 said bottleneck identification unit includes

means for, under a condition wherein the processing time is not within a predesignated range for the application servers, the number of which is equal to or smaller than a threshold value that is less than the  
25 application server count, identifying, based on monitoring results of processing time, a bottleneck in at least one of the plurality of application servers if the processing time

is not within a predesignated permissible range,

means for identifying a bottleneck in the database server if the processing time is not within the predesignated permissible range for the application servers, the number of which exceeds the threshold value;

wherein, if a bottleneck is identified in the application server, said load controller reduces multiplicity of the application program on the application server; and

wherein, if a bottleneck is identified in the database server, said load controller reduces multiplicity of the application program on the application servers.

3. An information system according to claim 1, wherein, among a plurality of types of application programs, for which priority levels are designated, that are to be executed by the application servers, each of the plurality of application servers executes the application program consonant with the type of transaction received from the terminal;

wherein, for each of the plurality of application servers, said process time monitoring unit monitors a processing time for the transaction by each of the application programs;

wherein, under a condition wherein the processing time for one of the application programs is not within a predesignated permissible range, said bottleneck identification unit identifies a bottleneck in the

application servers, based on the monitoring results for the processing time for each application program for each of the application servers; and

5 wherein said load controller reduces multiplicities with which the application servers wherein a bottleneck is identified executes the application program that has a priority level lower than the one application program.

4. An information system according to claim 1, wherein, among a plurality of types of application programs, for  
10 which priority levels are designated, that are to be executed by the application servers, each of the plurality of application servers executes the application program consonant with the type of transaction received from the terminal;

15 wherein, for each of the plurality of application servers, said process time monitoring unit monitors a processing time for the transaction by each of the application programs;

20 wherein, as for the application servers the number of which exceeds a predesignated threshold value, if the processing time for one of the application programs is not within a predesignated permissible range, said bottleneck identification unit identifies a bottleneck in the database server, based on the monitoring results for the processing  
25 time for each application program for each of the plurality of application servers; and

wherein said load controller reduces a priority

level, for the database server identified as a bottleneck, of a process in which the database is accessed based on a request by the application program that has a lower priority level than the one application program.

5        5.        An information system according to claim 1, wherein, among a plurality of types of application programs, for which priority levels are designated, that are to be executed by the application servers, each of the plurality of application servers executes the application program  
10        consonant with the type of transaction received from the terminal;

             wherein, for each of the plurality of application servers, said process time monitoring unit monitors a processing time for the transaction by each of the  
15        application programs;

             wherein, as for the application servers the number of which exceeds a predesignated threshold value, if the processing time for one of the application programs is not within a predesignated permissible range, said bottleneck  
20        identification unit identifies a bottleneck in the database server, based on the monitoring results for the processing time for each application program for each of the plurality of application servers; and

             wherein said load controller reduces a priority  
25        level, for the database server identified as a bottleneck, of a process by which the database is accessed based on a request by the one application program.

6. An information system according to claim 1, further comprising:

a process load monitoring unit, for monitoring, for each of the application servers, an amount consumed of an application server resource that is a resource owned by the application server and used by the application program;

wherein, if the processing time is not within a predesignated range for the application servers, the number of which is equal to or smaller than a threshold value that is less than the application server count, said bottleneck identification unit employs the monitoring results for the processing times required by the plurality of application servers to identify a bottleneck in the application servers,

wherein, under a condition wherein the amount consumed of the application server resource, for the application servers identified as a bottleneck, is not within a predesignated permissible range, said load controller reduces multiplicities with which the application servers identified as a bottleneck execute the application program.

7. An information system according to claim 6,

wherein said process load monitoring unit further monitors an amount consumed of a database server resource that is a resource owned by the database server and is to be used if the application program accesses the database;

wherein, under a condition wherein the processing time is not within a predesignated range for the

application servers, the number of which is greater than the threshold value, said bottleneck identification unit identifies a bottleneck in the database server, based on the monitoring results for the processing time required by each of the plurality of application servers; and

wherein, if the amount consumed of the database server resource, for the database server identified as a bottleneck, is not within the predesignated permissible range, said load controller reduces the multiplicities with which the plurality of application servers execute the application program.

8. An information system according to claim 1, further comprising:

a process load monitoring unit, for monitoring, for each of the plurality of application servers, amounts consumed of a plurality of application server resources, which are resources owned by the application server and are to be used by the application program, and amounts consumed of a plurality of database server resources, which are resources owned by the database server and are to be used if the application program accesses the database,

wherein said bottleneck identification unit employs the monitoring results for the processing time for each of the plurality of application servers to identify a bottleneck either in at least one of the plurality of application servers or in the database server, and

wherein said load controller includes

a load control information storage unit, for storing, in correlation with the plurality of application server resources of the application server and the plurality of database server resources of the database server, load control information that designates a load control process that said load controller should perform if a bottleneck is identified in a server that includes a pertinent resource, and if the amount consumed of the pertinent resource is not within a predesignated permissible range,

a load control information acquisition unit, for, if an amount consumed of a resource that is owned by a server wherein a bottleneck is identified is not within the predesignated permissible range, identifying the resource as a bottleneck, or for, if the application server resource is identified as a bottleneck, obtaining the load control information stored in the load control information storage unit, in correlation with the resource, or if the database server resource is identified as a bottleneck, obtaining the load control information stored in the load control storage unit, in correlation with the resource, and

a load control processor, for performing the load control process designated by the load control information in order to resolve the bottleneck of a server wherein the bottleneck is identified.

9. An information system according to claim 8, wherein, among the plurality of types of application

programs, each of the plurality of application servers performs the application program consonant with the type of the transaction received from the terminal;

wherein, for each of the plurality of application  
5 servers, said process time monitoring unit monitors a processing time for the transaction by each of the application programs;

wherein, said process load monitoring unit monitors, for the database server, the amounts consumed of a  
10 processor and a disk on which a database is stored, both of which are database server resources owned by the database server;

wherein, based on the monitoring results for the processing time required for each application program by  
15 each of the plurality of application servers, said bottleneck identification unit identifies a bottleneck in the database server, under a condition wherein the processing time required for the one application program is not within a predesignated range for the application  
20 servers, the number of which exceeds the threshold value;

wherein said load control information storage unit stores first load control information, which is correlated with a processor that is the application resource for the database server, and second load control information, which  
25 is correlated with a disk stored on the database;

wherein said load control information acquisition unit obtains the first load control information if the



amount consumed of the processor, owned by the database server where a bottleneck is identified, exceeds a predesignated amount consumed, or obtains the second load control information if the amount consumed of the disk, on  
5 which the database owned by the database server wherein a bottleneck is identified is stored, exceeds a predesignated amount consumed;

wherein, if the first load control information is received from said load control information acquisition  
10 unit, said load control processor reduces multiplicities for the plurality of application servers to execute the one application program;

wherein, if the second load control information is received from said load control information acquisition  
15 unit, said load control processor lowers a priority level for a process in which the database is to be accessed based on a request issued by an application program in a group that uses the disk designated by the control information for the database server and that has a lower priority level  
20 than the one application program.

10. An information system according to claim 8,

wherein one set of the load control information stored in the load control information storage unit is information for designating a plurality of the load control  
25 processes and of priority ranks for the load control processes;

wherein said load control processor performs one of

the load control processes designated by one set of the load control information, and if said load control processor detects that a bottleneck at a server where the bottleneck is identified has not been resolved during a predesignated setup time period, said load control processor performs another load control process at a priority rank lower than the load control process designated by the one set of load control information.

11. An information system according to claim 8,

wherein, of the plurality of types of application programs, each of the plurality of application servers performs the application program consonant with the type of transaction received from the terminal;

wherein, if the processing time for the one application program is not within a predesignated range for the application servers, the number of which is equal to or smaller than the threshold value that is less than the application server count, said bottleneck identification unit employs the monitoring results for the processing times required by the plurality of application servers, for each application program, to identify a bottleneck in the application servers;

wherein said load controller reduces a multiplicity for which the application servers, where the bottleneck is identified, execute the application program the priority level of which is set lower than the one application program; and

wherein, if detecting that the bottleneck in the application servers has not been resolved during the setup period, said load controller reduces further the multiplicity for the execution of the one application program by the application servers.

12. A load control method, for employing a computer to control an information system that includes a plurality of application servers, each of which can execute at least one application program for processing a transaction received from a terminal, and a database server, for accessing a database based on a request received from any one of the plurality of application servers, comprising the steps of:

permitting said computer to monitor processing time required for the application program to process a transaction received by an application server;

permitting said computer to identify, based on monitoring results of processing time, a bottleneck in at least one of the plurality of application servers for if the processing period time is not within a predesignated permissible range; and

permitting said computer to reduce a multiplicity of the application program on an application server identified as having a bottleneck.

13. A load control program, for employing a computer to control an information system that includes a plurality of application servers, each of which can execute at least one application program for processing a transaction received

from a terminal, and a database server, for accessing a database based on a request received from any one of the plurality of application servers, that permits said computer to function as:

- 5           a process time monitoring unit, for monitoring processing time required for the application program to process a transaction received by an application server;
- a bottleneck identification unit, for identifying, based on monitoring results of processing time, a
- 10   bottleneck in at least one of the plurality of application servers for if the processing period time is not within a predesignated permissible range; and
- a load controller, for reducing a multiplicity of the application program on an application server identified
- 15   as having a bottleneck.<sup>14</sup>     A recording medium on which a load control program according to claim 13 is recorded.